

МЕТОДЫ ОБРАБОТКИ И АНАЛИЗА СЛОЖНЫХ ПРИРОДНЫХ СИГНАЛОВ PROCESSING AND ANALYSIS METHODS OF COMPLEX NATURAL SIGNALS

Мандрикова О.В., Полозов Ю.А.

Институт космофизических исследований и распространения радиоволн ДВО РАН, ikir@ikir.ru

Natural signals have difficult multiscale structure, contain the local features arising during the casual moments of time and bearing the important information on process generating a signal. It essentially complicates process of their processing and does not allow to apply traditional methods of the analysis of time numbers. In the given work methods of allocation and classification of structural features of the natural signals, based on the wavelet-theory are offered. The given theoretical device allows to allocate local structures forming a signal that gives the chance to make their detailed analysis and identification in an automatic mode.

Регистрируемые природные сигналы являются нестационарными и имеют сложную многомасштабную внутреннюю структуру, содержат локальные особенности, имеющие априори неизвестную форму и длительность, возникающие в случайные моменты времени и несущие важную информацию о генерирующем сигнал процессе. Это существенно затрудняет процесс их обработки и не позволяет применять традиционные методы анализа временных рядов. В данной работе предлагаются методы выделения и классификации структурных особенностей природных сигналов, основанные на вейвлет-теории [1-3]. Данный теоретический аппарат позволяет в автоматическом режиме выделить формирующие сигнал локальные структуры, что дает возможность выполнить их детальный анализ и идентификацию. Предложенные методы имеют важное значение в геофизике, где весьма остро стоит задача выделения и идентификации аномальных явлений [3].

Идентификация регулярных составляющих сигнала производится с использованием конструкции вейвлет-пакетов [1-3] и нейронных сетей [4]. Используется многослойная сеть переменной структуры прямой передачи сигнала. Обучающая выборка сети формируется восстановленными вейвлет-компонентами. Построение структуры сети, ее обучение и оценка качества работы производятся совместно, что позволяет адаптировать этот процесс под внутреннюю структуру исходного входного сигнала.

Детальный анализ локальных структур сигнала основан на конструкции непрерывного вейвлет-разложения. Выполняется декомпозиция сложного сигнала на составляющие его компоненты и производится их анализ и классификация.

Предлагаемые методы опробованы на реальных геофизических данных, регистрируемых в Институте космофизических исследований и распространения радиоволн ДВО РАН. Полученные экспериментальные результаты исследований подтвердили их эффективность по сравнению с существующими аналогами. Используя данный теоретический аппарат, разработана система анализа и прогноза поведения важных характеристик геофизических сигналов.

Выделение регулярных составляющих сигнала

Особенностью аппарата нейронных сетей является необходимость правильного выбора входных данных. Они полностью определяют будущую структуру нейронной сети и позволяют оптимизировать процедуру ее обучения.

В процессе формирования обучающей выборки при моделировании сигналов необходимо учитывать следующие важные особенности: каждое значение зависит от предыдущих значений временного ряда; при обработке сигналов с шумом, он может исказить результат обучения сети.

Алгоритм формирования обучающей выборки. Массив данных $\{X(n)\}_{n=1}^M$ делится на L блоков длины T :

$$\{X(n)\}_{n=1}^N = \left(\{X(n)\}_{n=1}^T, \{X(n)\}_{n=2}^{T+1}, \dots, \{X(n)\}_{n=M-T}^M \right).$$

Производится отображение сигнала в пространство вейвлет-пакетов, в результате которого для каждого блока s получаем представление сигнала X в виде линейной комбинации разномасштабных составляющих:

$$X^s = X_1^s + X_2^s + \dots + X_p^s, \quad (1)$$

где $X_j^s \in W_j^{p_i}$, $W_j^{p_i}$ – пространства вейвлет-пакета.

Составляющие X_j^s в соотношении (1) единственным образом определяются последовательностями вейвлет-коэффициентов $\bar{c}_j^s = \left\{ c_{j,n}^s \right\}_{n=s, T+s-1}$, $c_{j,n}^s = \langle X^s, \Psi_{j,n} \rangle$. Каждая выделенная компонента X_j^s определяет подпространство $W_j^{p_i}$ пространства признаков сигнала X . Таким образом,

для каждого блока s получаем разнесение в пространстве признаков составляющих сигнала X . Все полученные компоненты (1) определяют пространство признаков исходного сигнала.

Возможен другой способ формирования блоков, например, с учетом определенных особенностей и свойств аппроксимируемых данных (сезонность, наличие пропусков и др.).

Составляющие, из которых сформируется обучающая выборка для нейронной сети, должны содержать характерные для сигнала признаки. Эти признаки определяют регулярные составляющие сигнала, для их идентификации построим нейронные сети для различных комбинаций входных образов (1) на основе следующего алгоритма:

Цикл 1: Для каждого блока данных s , используя составляющую X_1^s , выполняем вейвлет-восстановление сигнала (остальные составляющие при восстановлении заменяем нулями) и на основе комбинаций восстановленных данных из различных блоков формируем обучающее множество. Строим сеть 1, выполняем ее обучение и тестирование.

Цикл 2: Для каждого блока данных s , используя составляющие X_1^s, X_2^s , аналогично циклу 1, выполняем вейвлет-восстановление сигнала и на основе комбинаций восстановленных данных из различных блоков формируем обучающее множество. Строим сеть 2, выполняем ее обучение и тестирование.

И т.д.

Цикл p : Для каждого блока данных s , используя составляющие $X_1^s, X_2^s \dots X_p^s$, аналогично циклам 1, 2, выполняем вейвлет-восстановление сигнала и на основе комбинаций восстановленных данных из различных блоков формируем обучающее множество. Строим сеть p , выполняем ее обучение и тестирование.

Цикл $(p+1)$: Для каждого блока данных s , используя составляющие X_1^s, X_3^s , выполняем вейвлет-восстановление сигнала и на основе комбинаций восстановленных данных из различных блоков формируем обучающее множество. Строим сеть $p+1$, выполняем ее обучение и тестирование.

И т.д.

Таким образом, для каждой возможной комбинации имеющихся признаков строится нейронная сеть. На основе анализа результатов работы полученных нейронных сетей определяем «наилучшую»: в качестве «наилучшей» определяем сеть, имеющую наименьшую ошибку на тестовых данных

$$E_{N,\min} = \min_l \left(\frac{1}{D} \sum_{n=0}^{D-1} e_l(n)^2 \right), \text{ где } l = 1, \dots, r, \text{ } D - \text{длина тестовой выборки.}$$

Алгоритм построения структуры нейронной сети. Формирование структуры нейронной сети, при выполнении описанных выше циклов, производится на основе следующего алгоритма:

1. Строим сеть, состоящую из одного нейрона в первом слое и α нейронов во втором (выходном) слое. Производим ее обучение.
2. В первый слой сети добавляем один нейрон и вновь произведем ее обучение.
3. Аналогично нейроны продолжаем добавлять в первый слой, до тех пор, пока они улучшают общее качество аппроксимации сигнала сетью.
4. Если очередной нейрон не дает улучшения характеристик сети, удаляем его, а в сеть добавляем промежуточный слой, состоящий из одного нейрона. Аналогичным образом производим обучение сети, и если добавленный слой повысил уровень аппроксимации, то в него начинаем добавлять нейроны по принципу п.2, 3.
5. Подобным образом создается столько слоев, сколько необходимо для достижения требуемого качества работы сети при заданном массиве входных данных.

Выделенные компоненты сигнала аппроксимируются нейронной сетью и являются регулярными. Другие его составляющие из соотношения (1) содержат определенную долю информации о природном явлении, они будут использованы при дальнейшем анализе. Это обеспечивает возможность полного сохранения информации о природном процессе.

Детальный анализ и классификация локальных структур сигнала

Так как вейвлет Ψ имеет нулевое среднее значение, непрерывное вейвлет-преобразование

$$WX(a, b) = \int X(t) \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) dt$$

измеряет изменение X в окрестности точки b , размер которой

пропорционален масштабу a . При стремлении масштаба a к нулю вейвлет-коэффициенты характеризуют локальные свойства функции X в окрестности точки b .

В работах [1, 2] доказано, что убывание амплитуды коэффициентов вейвлет-преобразования $|WX(a, b)|$ в зависимости от масштаба a связано с равномерной и точечной гладкостью Липшица сигнала. Когда масштаб a убывает, амплитуды вейвлет-коэффициентов $|WX(a, b)|$ имеют быстрое убывание до нуля в областях, где сигнал не содержит локальных особенностей. Основываясь на данной теории можно считать, что функция X в окрестности точки v имеет локальную особенность, если $|WX(a, b)|$ не удовлетворяет условию [2]

$$|WX(a, b)| \leq A' a^{\alpha+1/2}, \quad (2)$$

A' - некоторая константа; α - показатель Липшица, что тождественно равномерному условию Липшица.

Характеризовать X в точке b бывает достаточно трудно, потому что X может иметь различные виды особенностей, присутствующих в окрестностях точки b . Более ясную интерпретацию здесь дает **конус влияния точки v** : если вейвлет Ψ имеет компактный носитель, равный $[-C, C]$, то множество точек (a, b) таких, что точка v содержится в носителе $\Psi_{a,b}(t) = a^{-1/2}\Psi((t-b)/a)$, определяют **конус влияния точки v** масштабно-пространственной плоскости. Так как носитель $\Psi((t-b)/a)$ равен $[b - Ca, b + Ca]$, то конус влияния v определяется неравенством

$$|b - v| \leq Ca \quad (3)$$

Если b находится в конусе влияния v , то $WX(a, b) = \langle X, \Psi_{a,b} \rangle$ зависит от значения X в окрестности v .

Рассмотрим теперь **дискретный сигнал с шумом**:

$$X[n] = f[n] + e[n], \quad (4)$$

где $e[n]$ - шум.

На основе конструкции дискретного вейвлет-разложения сигнала получаем:

$$X(t) = \sum_{j,n \in \mathbb{Z}} d_{j,n} \Psi_n^j(t), \quad (5)$$

где $d_{j,n} = \langle X, \Psi_{j,n} \rangle$. Эти бесконечные ряды называются *вейвлет-рядами* и сходятся в $L^2(\mathbb{R})$.

Здесь каждая $\Psi_{j,n}$ определяет свой частотно-временной прямоугольник в частотно-временной плоскости. Сдвиги $\Psi_{j,n}$ на $2^j n$, с временным n и частотным 2^j параметрами, покрывают эту плоскость прямоугольниками минимальной площади, как обусловлено принципом неопределенности. В этом случае процедура выделения локальных особенностей функции X может быть построена на анализе детализирующих коэффициентов (см. (5)) $\bar{d}^j = \{d_{j,n}\}_{n \in \mathbb{Z}}$, которые для малых j определяют мелкомасштабные компоненты сигнала. Из соотношения (2) следует, что данная процедура может быть легко реализована на **основе применения пороговой функции**:

$$f = \sum_{j,n} P_T(d_{j,n}) \Psi_{j,n},$$

$$P_T(y) = \begin{cases} y, & \text{если } |y| \geq T, \\ 0, & \text{если } |y| < T, \end{cases}$$

где T – порог. В этом случае встает вопрос определения порога T . Имея в наличии разномасштабные составляющие сигнала, порог необходимо определять для каждого масштабного уровня j .

Поскольку в момент возникновения **аномального явления** в сигнале возникает особенность, для ее выделения может быть применена пороговая функция. При этом выбор порога для масштаба j может быть основан на следующей процедуре:

$$T_j = \min |d_{I_{j,n}}|,$$

где множество индексов $I_{j,n} = \{(j, n) : |n - v'| \leq C_j' j\}$ (см. (3)), v' - момент возникновения аномального явления, C_j' - параметр, определяющий размер интервала длительности аномалии на масштабе j .

Из равенства (4) следует, что некоторые коэффициенты вейвлет-преобразования $d_{j,n}$ являются шумовыми. В работе [5] доказано, что они могут быть выделены на основе применения пороговой функции, где порог $T_e = \sigma^2$, $\sigma^2 \approx \text{Med} \left(\left\langle X, \Psi_{j,n} \right\rangle \right)_{0 \leq n < \frac{k}{2}}$, где Med – медиана, j – наименьший масштаб, k – длина сигнала. Поскольку порог в данном случае больше максимального уровня коэффициентов шума: $T_j > T_e$, данная процедура обеспечит подавление шумовой составляющей сигнала.

Интервал длительности аномалии и масштабы j служат характеристиками происходящего физического процесса. Имея информацию о моменте возникновения особенности, ее длительности и масштабах, можно выполнить следующую классификацию: (а) длительность аномалии C'_{j_i} , где j_i – наибольший масштаб, длительность аномалии на масштабе j : C'_j ; (б) масштабные уровни аномалии: $j \in I_{j,n}$.

В данной работе рассматривались сигналы критической частоты ионосферы. Они имеют сложную структуру, различную амплитуду колебаний в разные сезоны (4-7 МГц – летом и 2-7 МГц – зимой), содержат суточный ход и переходные процессы. Поэтому пороговое значение T_j определялось в пределах скользящего временного окна, равного 720 отсчетам сигнала (что соответствует временному интервалу в один месяц), следующим образом: $T_j = \text{med}_n(d_{j,n}) + 2,5St_j$, где

$$St_j = \sqrt{\frac{1}{N-1} \sum_{n=1}^N (d_{j,n} - \bar{d}_{j,n})^2}, \quad \bar{d}_{j,n} - \text{среднее значение, } \text{med} - \text{медиана, которая характеризует средний}$$

уровень процесса. Коэффициент 2.5 определялся статистически. На рис. 1 показан результат обработки сигнала критической частоты за январь-февраль 1986 года. Моменты возникновения аномалий в сигнале выделены на рис. 1.а красным цветом. Стрелкой отмечен момент землетрясения (2 марта 1986 г., $E = 12,8$). Анализ рис. 1.а показывает, что накануне землетрясения в сигнале выявлена многомасштабная аномалия, которая возникла за 23 дня до события и проявила себя на масштабных уровнях 5-15 и 23-38 (что соответствует периодам колебаний 6-30 часов). Длительность аномалии составляла 6 часов.

На рис. 1.б показан результат работы нейронной сети при подаче на ее вход восстановленной компоненты X_1^S сигнала критической частоты f0F2. Исследование и анализ сигналов f0F2 показали, что данная компонента является регулярной. Анализ графика на рис. 1.б показывает, что в моменты возникновения аномалий происходит значительное увеличение дисперсии ошибки сети. Полученные результаты подтверждают эффективность предложенных методов.

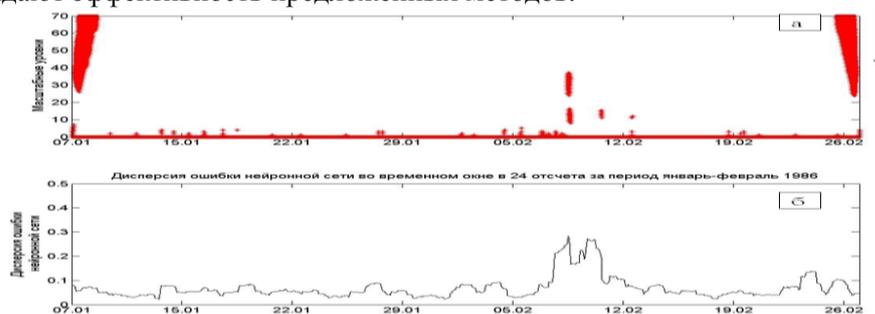


Рис. 1. Выявление аномалий в сигнале критической частоты f0F2 на основе предложенных методов. а – выявление аномалии по порогу, б – дисперсия ошибки нейронной сети, рассчитанная во временном окне, равном 24 отсчета. Стрелкой отмечено произошедшее землетрясение.

Выводы. В работе предложены методы выделения структур сложных природных сигналов. Данные методы позволяют выделить регулярные составляющие, локальные особенности и подавить шум. Проведенная апробация методов на природном сигнале критической частоты f0F2 показала их эффективность.

Литература

1. Ingrid Daubechies, Ten Lectures on Wavelets. Пер. с английского, – Ижевск: НИЦ «Регулярная и хаотическая динамика».– 2001.
2. Stephane Mallat, A Wavelet tour of signal processing. Пер. с английского, – М.: Мир, 2005.
3. Богданов В.В., Геппенер В.В., Мандрикова О.В. Моделирование нестационарных временных рядов геофизических параметров со сложной структурой. – С.-Петербург: ЛЭТИ. 2006. – 107с.

4. Галушкин А.И. общая редакция. Нейроматематика. Кн.6. М.: ИПРЖР. 2002. 448 с.
5. Мандрикова О.В., Горева Т.С. Метод идентификации структурных компонентов сложного природного сигнала на основе вейвлет-пакетов. - Научно-технический журнал "Цифровая обработка сигналов" № 1-2010 г., с.45-50.

МАГНИТНОЕ ПОЛЕ БАЙКАЛЬСКОЙ РИФТОВОЙ ЗОНЫ ПО ДАННЫМ ВЫСОКОТОЧНОЙ НАЗЕМНОЙ СЪЕМКИ

THE MAGNETIC FIELD OF BAIKAL RIFT AS RESULTS OF PRECISION SURFACE SURVEY

Р.А.Рахматулин¹, С.Ю.Хомутов², В.В.Харченко¹, Ю.В.Липко¹

¹Институт солнечно-земной физики СО РАН, rav@iszf.irk.ru

²Алтае-Саянский филиал Геофизической службы СО РАН, khomutov@gs.nsc.ru

The results of the precision magnetic survey during 2009-2010 on Lake Baikal ice to the north of the Olkhon Island and on the Olkhon Island are presented. The main traverse of survey was routed along the Academic ridge from Olkhon Island to Bolshoy Ushkany Island and other tracks were routed across the main track. The Overhauser proton magnetometer POS-1 used for measurements of total field F . The angular components of the total vector (declination D and inclination I) measured by DI-fluxgate 3T2KP and the spatial variations of the vertical component dZ measured by magnetometer M-27M. The spatial resolution along the route is about 5 km for F and dZ and about 10 km for total field. The accuracy of the measurement is estimated as about 0.5 nT, 1 arcmin and 0.2 arcmin for F , D and I , accordingly. The data from magnetic observatories Uzur (Olkhon Island), "Patrony" (Irkutsk) and "Klyuchi" (Novosibirsk) were used for control the geomagnetic variations during survey. As the results of magnetic survey the maps of distribution of magnetic field in central part of Baikal Lake between Olkhon and Bolshoy Ushkany islands were obtained. This maps were compared with the results of previous surveys. The large-scale magnetic anomaly which not presented on magnetic maps of this region was found in survey area. The features of the precise magnetic survey on ice surface are discussed.

Экспериментальные исследования Байкальской рифтовой зоны, включая ее центральную часть – оз.Байкал, имеют длительную историю. В тоже время, магнитное поле этого региона изучено недостаточно подробно, преимущественное по результатам аэромагнитной съемки в 60-х годах [1], гидромагнитной съемки на акватории в начале 70-х годов [2] и по данным непрерывных измерений в отдельных пунктах [3]. В связи с этим ИСЗФ СО РАН были инициированы работы по проведению высокоточной магнитной съемки полного вектора магнитной индукции на льду центральной части оз.Байкал, к северу от о.Ольхон, на котором расположена обсерватория "Узур".

Используемая аппаратура, район работ, условия измерений

Комплекс для проведения компонентной магнитной съемки с поверхности льда включал высокоточные магнитометры обсерваторского класса, в том числе:

- 1) протонный оверхаузеровский магнитометр POS-1 (УГТУ-УПИ, г.Екатеринбург), обеспечивающий измерения модуля вектора магнитной индукции F с погрешностью не хуже 0.2 нТл. Для управления и регистрации данных использовался полевой накопитель DLPOS;
- 2) DI-магнитометр для измерения магнитного склонения D и наклона I на базе немагнитного теодолита 3T2KP и феррозондового датчика в качестве нуль-индикатора положения. Инструментальная погрешность угловых параметров магнитного поля оценивается не более 5-10". Для установки прибора в пункте измерения использовался специальный немагнитный штатив;
- 3) пешеходный магнитометр M-27M для измерения пространственных изменений вертикальной компоненты ΔZ с приборной погрешностью 4-5 нТл.

Все приборы были протестированы на обсерваториях "Ключи" (Новосибирск) и "Патроны" (Иркутск), в том числе и в зимних условиях. Геодезические координаты и точное время в полевых условиях определялись с помощью GPS-навигатора Garmin. Для перемещения вдоль маршрута съемки использовались два автомобиля УАЗ-3909.

Район работ предварительно был выбран по результатам анализа магнитной изученности оз.Байкал, с учетом наличия стационарных наблюдений (станция "Узур" на севере о.Ольхон) и возможностью обеспечить приемлемые условия для измерений и ночевки. Первый профиль для съемки был запланирован вдоль Академического хребта от о.Ольхон до о.Большой Ушканий, два других – в поперечном направлении. Положение района проведения магнитной съемки показано на рис.1. Общая протяженность профилей около 150 км, количество пунктов dZ, F -измерений – 32, пунктов измерений полного вектора (D, I, F) – 16.